

## EXPERT REPORT

**Michael R. Smith, J.D., Ph.D.**

The conclusions reached by Dr. Eli Coston in their report entitled “Traffic Stops by the Richmond Police Department: July 1, 2020 – December 6, 2020” are unsupported by the data and analysis they conducted and rely upon elementary statistical techniques that no experienced social scientist familiar with racial disparity research in policing would rely on to draw inferences of bias. Below, I will address each of the primary claims that Dr. Coston makes in the concluding section of their report and demonstrate why they cannot and should not make them based on the analyses they conducted:

- Black drivers are at a significant disadvantage compared to White drivers in Richmond
  - Black drivers are more likely to experience arrests than White drivers, particularly after high discretion traffic stops
  - Black drivers are more likely to experience searches than White drivers
- Spatial analyses of stops show patterns “implicating race as a factor in traffic stops”

Below, I discuss each of these claims, but I take them in reverse order. I discuss first the initial decision to stop a driver and why the spatial analyses Dr. Coston conducted cannot be used to draw valid conclusions about racial disparities among Black and White drivers, let alone support an inference of bias. Next, I address the claim of more punitive outcomes (i.e., searches and arrests) allegedly experienced by Black drivers *after* a stop has occurred and why those claims, too, are scientifically invalid.

### **The Initial Stop Decision**

For more than 20 years, researchers have studied whether the race and/or ethnicity of drivers influences the decisions by police officers to make traffic stops. In 2001, a colleague and I published the first peer-reviewed article to investigate racial disparities in traffic stops and stop outcomes using traffic stop data from Richmond (Smith & Petrocelli, 2001), a paper that has since been cited more than 270 times by other scholars. In those days, there was little scholarship on “racial profiling,” and there was no methodological roadmap for scholars to follow. Even then, though, we recognized that examining the raw percentages of drivers stopped by race is meaningless unless, at a minimum, those percentages can be compared against an estimate of the population of *drivers* in the jurisdiction of interest (Alpert et al., 2004; Fridell, 2004; Ridgeway & MacDonald, 2010). In the years following the publication of our paper, a robust literature has developed on racial/ethnic disparities in traffic stops, and science has evolved and improved on the key issue of *benchmarks* (Alpert et al., 2007; Grogger & Ridgeway, 2006; McLean & Rojek, 2016; Ritter, 2017; Smith et al., 2021; Tillyer et al., 2010).

Today, census-based benchmarking is no longer accepted as a scientifically valid technique for comparing against police traffic stop data. First generation studies, including our original Richmond study, used census or age-adjusted (aged 15-16 and older) census data to compare against the racial composition of drivers stopped by the police (Smith et al., 2021). Census data is free and readily available and makes a convenient benchmark. Unfortunately, as later comparative studies made clear (see Alpert et al., 2004), the population of persons who *live* in an area often serves as a poor representation of persons who *drive* in an area or who are *at risk* for being stopped by the police. Census benchmarking does not account for out-of-area drivers, differential exposure to the police due to differences in police deployment patterns within a city or state, or differences in driving behavior across racial groups (Ridgeway & MacDonald, 2010; Smith et al., 2019; Tillyer et al., 2010), among other factors. While all benchmarks suffer from certain limitations, census-based benchmarks are so far off the mark as valid estimates of the actual driving and/or traffic violating populations that informed social scientists should no longer use them (Ridgeway & MacDonald; Smith et al., 2021).

Most social scientists who regularly study racial disparities in traffic stops would agree that direct field observation of drivers is the “gold standard” for benchmarking (McLean & Rojek, 2016; Alpert et al., 2004). Unfortunately, field observation is expensive and time-consuming, and due to those limitations, it has been used in only limited areas (e.g. at selected intersections or stretches of highway) and in a relatively few studies (Alpert et al., 2004; Lange et al., 2001; Zingraff et al., 2000). Instead, two alternative benchmarks have emerged that have been tested in multiple cities and now appear regularly in the peer-reviewed literature.

Grogger & Ridgeway (2006) developed a widely used statistical test, which does not rely upon external benchmark data, to compare rates at which minority and White drivers are stopped by the police. Often referred to as the “veil of darkness” (VOD) technique, it makes use of the natural variation in daylight that occurs across the year and with changes in daylight savings time to compare the proportion of drivers stopped at night to those stopped during the day by racial group. Theoretically, if more minority drivers are stopped during daylight hours when police can more easily ascertain their race, then this provides evidence of possible racial bias, particularly if there are no differences in the rates at which White drivers are stopped during the day compared to at night. This technique has been widely replicated and reported in the scientific literature (Channin, et al., 2016; Pierson et al., 2020; Ritter & Bael, 2009; COPS, 2016; Ross et al, 2016; Taniguchi et al., 2016; Vito et al., 2020; Worden et al., 2012).

In addition, Alpert and colleagues (including this author) pioneered a benchmarking technique using not-at-fault drivers in two vehicle crashes as a proxy for the driving population (Alpert et al., 2004; Lovrich et al., 2007). Drawing upon the traffic safety literature that has long-used crash data for an unbiased estimate of risk among subpopulations (usually age and gender-related), Alpert and colleagues theorized that the racial composition of not-at-fault drivers would serve as

an unbiased estimate of persons *driving* (as opposed to residing) in a jurisdiction that could be compared against the percentage of drivers of different races stopped by the police. This technique has since been extended to the use of *at-fault* drivers as a proxy for traffic law *violators* (Withrow, 2015) and was recently tested against VOD in light of recent critiques of both VOD and crash-derived benchmarks (Smith et al., 2021). At the end of the day, while benchmarking is an inexact science, both traffic crash benchmarks and the VOD approach offer significant improvements over census-derived benchmarks, which no longer represent scientifically acceptable estimates of drivers at risk for being stopped by the police.

### **Coston's Use of Census-Derived Mapping to Make Inferences of Bias**

Dr. Coston's use of maps to draw qualitative inferences about racial bias in stops made by the Richmond Police Department is scientifically invalid. Mapping in general, and cluster and/or heat mapping in particular, can be a useful technique for visualizing the spatial distribution of events (crimes, arrests, traffic stops) within a city. It cannot be used to draw causal inferences about the influence of race on traffic stops as Dr. Coston attempts to do.

First, the clustering of traffic stops by race can be influenced by a variety of factors. From a motorist perspective, the risk for being stopped by the police includes factors such as:

- Where people drive
- When they drive
- How often they drive
- What they drive (type of vehicle and its condition)
- How they drive (risky or traffic law violating driving behavior)
- Who they are (i.e. race, ethnicity, age, gender, etc.) (Tillyer et al., 2010)

From the police perspective, possible mechanisms that can influence stop disparities include overt bias or racial animus, unconscious or implicit bias, police deployment patterns or a mixture of all three (Tomaskovic-Devey et al., 2004). Police typically deploy more officers to neighborhoods with higher crime rates and which generate more calls for service (Engel, Smith, & Cullen, 2012; Parker, MacDonald, Alpert, Smith, & Piquero, 2004; National Research Council, 2004). Moreover, modern police crime control strategies increasingly rely on the deployment of officers to crime "hot spots" where crime mapping indicates offenses have recently occurred and are expected or predicted to occur again (Braga et al, 2019). As Tomaskovic-Devey et al. (2004) note, police deployment variables often correlate with race and ethnicity in America's urban neighborhoods.

Simple mapping techniques like those used by Dr. Coston show only *where* people drive and *who they are* (race) but account for none of the factors that otherwise influence the racial composition of traffic stops or how or why they occur in some areas more than others. The final heat map overlay of stops of Black drivers over the racial composition of Richmond neighborhoods (Coston, Fig. 5) is merely a rudimentary type of census-based benchmarking and suffers from all of the weaknesses of census benchmarks discussed above. It is rudimentary

because the definition of what constitutes “predominate” populations of various racial or ethnic groups is unspecified, and the resolution of the map and the boundaries between “predominate” racial and ethnic neighborhoods is so diffuse as to be meaningless. Moreover, because residential populations do not reflect *racial driving* patterns or the roadway features, travel patterns, or social and economic factors that may influence who actually *drives* in a jurisdiction, they cannot be used to draw inferences of racial bias simply because clusters of traffic stops involving Black drivers occur in ill-defined transitional areas of the city. One might logically expect, for example, that clusters of stops of Black drivers might occur in such transitional zones as Black drivers commute from where they may live to where they may work along major thoroughfares.

Dr. Coston relies on outmoded and rudimentary analytic techniques to draw invalid inferences about racial bias in traffic stops. Without paying careful methodological attention to the comparison of stops of Black drivers to scientifically-accepted estimates of the racial composition of *drivers* within those transitional zones, no valid inferences about racial bias in traffic stops can be made.

In sum, I have never seen an analysis like this in a peer-reviewed scientific publication, and for good reason. The scientific community of scholars who study racial disparities in traffic stops would not accept this type of analysis to support an inference of racial bias in traffic stops. It is far too imprecise and explicitly makes use of outmoded residential census-based estimates to reach invalid and unsupportable conclusions about the influence of race on the decision by police officers to stop drivers in Richmond, Virginia.

### **Post-Stop Outcomes**

Like traffic stops themselves, police activities that occur *after or during* a stop have been the subject of a great deal of research over the past 20 years. Outcomes such as citations, warnings, searches, and arrests all have been studied extensively in the literature, and there is some empirical consistency across studies in how driver race/ethnicity can impact those outcomes *after controlling for* other factors also known to correlate with them (Alpert et al., 2006; Alpert et al., 2007; Engel et al., 2012; Ridgeway, 2009; Smith & Petrocelli, 2001; Smith et al., 2017).

Key to understanding the influence of driver race/ethnicity on post-stop outcomes is to appropriately account or statistically control for factors that may confound the relationship between driver race and outcomes such as searches, arrests, or citations. If the question is whether drivers of a certain minority group are arrested more frequently than Whites, then it is important to control for the severity of the violation that led to the stop as well as the seriousness of the offense for which the arrest occurred. For example, intoxicated drivers are more likely to be arrested than speeders. Likewise, drivers who have pre-existing warrants on file with NCIC

will almost always be arrested, whereas police have greater discretion over on-view arrests based on probable cause developed during the stops themselves (Alpert et al, 2006).

Before valid inferences about bias can be made based on observed disparities in post-stop outcomes, researchers must *control for* the influence of exogenous factors such as pre-existing warrants or the seriousness of the offense on the arrest outcome being modeled. This is typically done using multivariate regression, which is a statistical technique used to isolate the influence of one variable on the outcome of interest while controlling for other known influences that may themselves be correlated with the outcome (Mertler & Reinhart, 2017). For searches, researchers must control for (or eliminate from consideration) *low discretion* arrests by controlling for the reason for the search. Searches incident to arrest or inventory searches are considered low discretion searches and are often required by agency policy or for officer safety reasons. Experienced analysts recognize that valid inferences about the influence of race on searches cannot be made unless the analyst controls for low versus high discretion search rationales (Alpert et al., 2006; Alpert et al., 2007; Smith et al., 2017).

### **Coston's Inferences of Bias in Post-Stop Outcomes**

Dr. Coston's conclusions that Black drivers were more likely to experience searches or arrests than Whites based on simple *bivariate* analyses without controlling for other variables known to correlate with those outcomes such as warrants, reasons for the search, offense severity, or driver behavior or demeanor (Engel et al., 2011) are not scientifically supportable. Causal inferences are simply not possible with non-parametric tests like the chi-square statistic and its companion test for relationship strength – Cramer's V. Chi-square can only tell us that two variables are related to one another; it cannot be used to conclude that one variable has any causal effect on the other (Imbens & Rubin, 2015).

In order to infer that one variable (race) *caused* another (search or arrest) to occur, three conditions must be met:

- The variables must be associated or *correlated* with one another
- The cause must have *preceded* the effect in time (it must have come first)
- The effect must not have been caused by some other *unmeasured* factor (non-spuriousness)

In addition, research methodologists note the importance of two other conditions in specifying cause and effect:

- There must be a reasonable *mechanism* to explain how one thing *caused* another to occur
- The *context* of the observed relationship should be specified and accounted for (Bachman & Schutt, 2007)

To help meet the rigorous requirements for *causal inference* social scientists attempt to design research studies that, whenever possible, account for all five of these conditions. In the case of

racial disparity research on police traffic stops and their outcomes, social scientists often turn to multivariate regression (discussed above) to help control for as many relevant factors as possible that might have influenced the observed outcome such as a citation, search, or arrest. In addition to driver race and ethnicity, important factors to take into consideration in a multivariate statistical model include things such as the reason for the stop, the seriousness of the traffic infraction or criminal offense for which the stop was made, the reason for the search, the time and location of the stop, the actions and demeanor of the driver, the condition of the vehicle itself, pre-existing warrants, and the environmental context of the area where the stop occurred, especially its crime rate or the type of roadway where the stop took place.

The literature is replete with studies that carefully measured as many of these variables as possible and controlled for them in multivariate statistical models that far exceeded the explanatory power of the rudimentary statistical techniques used by Dr. Coston (Alpert et al., 2006; Engel et al., 2012; Ridgeway, 2009; Smith & Petrocelli, 2001; Smith et al., 2017). Yet, even in those studies, the authors were careful not to infer *racial bias* as the cause of any observed disparities because of their inability to rule out the influence of other variables unavailable to them in their analyses.

A 2015 case from the Middle District of North Carolina is instructive on the limits of even carefully constructed multivariate regression models in proving racial bias in police traffic stop cases (*United States v. Johnson*, 2015). In *Johnson*, the U.S. government alleged that the sheriff of Alamance County engaged in a pattern or practice of discriminatory law enforcement in violation of the Constitution and the Violent Crime Control and Law Enforcement Act of 1994 (42 U.S.C. § 14141). Specifically, the government alleged that Sheriff Johnson's office (through its deputies) subjected Hispanic motorists to unreasonable searches, arrests, and traffic checkpoints in violation of their constitutional rights.

To help prove its case that Hispanics were discriminated against in searches and arrests in Alamance County, the government retained a highly qualified criminologist from the University of Pennsylvania (Dr. John MacDonald) who carried out a multivariate regression analysis predicting stop outcomes based on ethnicity of the drivers. Dr. MacDonald controlled for the generic reason for the stops available from the official data he analyzed (e.g. "speed limit violation" or "vehicle equipment violation"), but he was unable to control for the *seriousness* of the driver's conduct, which might have accounted for the dissimilar post-stop arrest and citation outcomes found between Hispanic and non-Hispanic drivers. The court concluded that Dr. MacDonald's statistical evidence was insufficient to prove that Hispanics were similarly situated to non-Hispanics with respect to the post-stop outcomes they received because another highly relevant variable – the reason for the arrest or citation – was not included in his analysis.

Similarly, with respect to his search analyses, Dr. MacDonald was unable to control for the *reason for the search* when he found that Hispanic drivers were more likely to be searched but



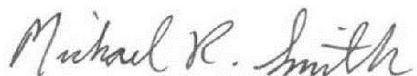
less likely to be found with contraband compared to non-Hispanic drivers. The court noted the importance of controlling for non-discretionary searches, such as those conducted incident to arrest, and concluded that the government had failed to prove that ethnicity (Hispanic heritage) was the *cause* of the disparate search and hit rates observed among Hispanic drivers.

Thus, even *multivariate regression* models much stronger than the simple bivariate tests used by Dr. Coston often cannot be used to infer causal relationships because they cannot demonstrate that the group receiving the harsher outcome (e.g., search or arrest) was similarly-situated to the group not receiving those outcomes. As a result, careful social scientists do not infer bias based merely on observed differences in the treatment by police of persons from different racial groups. Such differential treatment may have nothing to do with race but instead may reflect *legitimate* differences between racial groups. For example, if one group is more likely than another to have outstanding arrest warrants on file, then members of that group likely would experience higher *arrest rates* than persons from the other group. Likewise, if one group drove recklessly or under the influence of alcohol or drugs more often than another, then this might explain higher observed arrest rates for that group.

Similarly, observed differences in searches between groups might be explained by higher or lower *contraband carry rates*, *higher rates of arrest* (and searches incident to those arrests), or *higher rates of vehicle impoundment* (and any resulting, *non-discretionary* inventory searches) among some racial groups compared to others. Accounting for the underlying *reason* for a search is critical to any search disparity analysis, but Dr. Coston's rudimentary bivariate analysis did not take this critical variable into account. Based on the analysis conducted, Dr. Coston cannot, with legitimate scientific rigor, conclude that the race of drivers in Richmond has a *causal influence* on the whether they are stopped, searched, arrested, or cited by officers from the Richmond Police Department.

### Conclusion

Dr. Coston's statistical and mapping analyses cannot support an inference of racial bias in traffic stops or post-stop outcomes received by Black drivers in Richmond. They fail to account for numerous non-racial factors known in the scientific community to correlate with stops and post-stop outcomes such as searches, arrests, and citations, and they do not meet the scientific threshold for causal inference.



Michael R. Smith, J.D., Ph.D.

## References

- Alpert, G.P., Becker, E., Gustafson, M.A., Meister, A.P., Smith, M.R., Strombom, B.A. (2006, February). *Pedestrian and motor vehicle post-stop data analysis report*. Los Angeles California: Analysis Group.  
[https://www.analysisgroup.com/globalassets/uploadedfiles/content/insights/cases/lapd\\_data\\_analysis\\_report\\_07-5-06.pdf](https://www.analysisgroup.com/globalassets/uploadedfiles/content/insights/cases/lapd_data_analysis_report_07-5-06.pdf)
- Alpert, G.P., Dunham, R., & Smith, M.R. (2007). Investigating racial profiling by the Miami-Dade Police Department: A multimethod approach. *Criminology & Public Policy*, 6(1), 25-56.
- Alpert, G. P., Smith, M. R., & Dunham, R. G. (2004). Toward a better benchmark: Assessing the utility of not-at-fault traffic crash data in racial profiling research. *Justice Research and Policy*, 6(1), 43-69.
- Bachman, R. & Schutt, R.K. (2007). *The practice of research in criminology & criminal justice* (3<sup>rd</sup> ed.). Thousand Oaks, CA: Sage.
- Channin, J., Welsh, M., Nurge, D., & Henry, S. (2016). Traffic enforcement in San Diego, California: An analysis of SDPD vehicle stops in 2014 and 2015. San Diego, CA: San Diego State University.  
<https://www.sandiego.gov/sites/default/files/sdpdvehiclestopsfinal.pdf>
- COPS. (2016). *Collaborative reform initiative: An assessment of the San Francisco police department*. Washington, DC: Office of Community Oriented Policing Services.  
[https://www.sanfranciscopolice.org/sites/default/files/2021-04/DOJ\\_COPS%20CRI\\_SFDPD%20OCT%202016%20Assessment\\_0.pdf](https://www.sanfranciscopolice.org/sites/default/files/2021-04/DOJ_COPS%20CRI_SFDPD%20OCT%202016%20Assessment_0.pdf).
- Engel, R. S., Cherkauskas, J. C., Smith, M. R., Clutter, J., & Grothoff, G. (2012) Traffic stop data analysis study: Year 5 (2010) final report. *Institute of Crime Science, University of Cincinnati*.
- Engel, R. S., Smith, M. R., & Cullen, F. T. (2012). Race, place, and drug enforcement: Reconsidering the impact of citizen complaints and crime rates on drug arrests. *Criminology & Public Policy*, 11, 603–635.
- Engel, R.S., Tillyer, R., Klahm, C., & Frank, J. (2011). From the officer's perspective: A multilevel examination of citizens' demeanor during traffic stops. *Justice Quarterly*, 29, 650-683.



- Fridell, L. A. (2004). *By the numbers: A guide for analyzing race data from vehicle stops*. Washington, D.C.: Police Executive Research Forum.
- Grogger, J., & Ridgeway, G. (2006). Testing for racial profiling in traffic stops from behind a veil of darkness. *Journal of the American Statistical Association*, 101, 878-887.
- Imbens, G.W. & Rubin, D.B. (2015). *Causal inferences for statistics, social, and biomedical sciences: An introduction*. New York: Cambridge University Press.
- Lange, J. E., Blackman, K. O., Johnson, M. B. (2001). *Speed violation survey of the New Jersey Turnpike: Final report*. Calverton, MD: Public Services Research Institute.
- Lovrich, N. P., Gaffney, M. J., Mosher, C. C., Pratt, T. C., & Pickerill, M. J. (2007). *Results of the monitoring of WSP traffic stops for biased policing*. Pullman, WA: Division of Governmental Studies and Services, Washington State University.
- McLean, K. & Rojek, J. (2016). Traffic stops, race, and measurement. In B.M. Huebner & T.S. Bynum (Eds.), *The handbook of measurement issues in criminology and criminal justice* (pp. 452-472). Hoboken, NJ: Wiley.
- Mertler, C.A. & Reinhart, R.V. (2017). *Advanced and multivariate statistical methods* (6<sup>th</sup> ed.). New York: Routledge.
- National Research Council. (2004). *Fairness and effectiveness in policing*. Washington, D.C.: National Academies Press.
- Parker, K. F., MacDonald, J. M., Alpert, G. P., Smith, M. R., & Piquero, A. R. (2004). A contextual study of racial profiling: Assessing the theoretical rationale for the study of racial profiling at the local level. *American Behavioral Scientist*, 47, 943–962.
- Pierson, E., Simoiu, C., Overgoor, J., Corbett-Davies, S., Jenson, D., Shoemaker, A., Ramachandran, V., Barhouty, P., Phillips, C., Shroff, R., & Goel, S. (2020). A large-scale analysis of racial disparities in police stops across the United States. *Nature Human Behavior*, 4, 736-745. <https://doi.org/10.1038/s41562-020-0858-1>.
- Ridgeway, G. (2009). Cincinnati Police Department traffic stops: Applying RAND's framework to analyze racial disparities. RAND: Center on Quality Policing. <https://www.rand.org/pubs/monographs/MG914.html>.

- Ridgeway, G., & MacDonald, J. (2010). Methods for assessing racially biased policing. S.K. Rice & M.D. White (Eds.), *Race, ethnicity, and policing: New and essential readings*, (pp. 180-204). New York: NYU Press.
- Ritter, J.A. (2017). How do police use race in traffics stops and searches? Tests based on observability of race. *Journal of Economic Behavior & Organization*, 135, 82-98.
- Ritter, J. A., & Bael, D. (2009). Detecting racial profiling in Minneapolis traffic stops: A new approach. Minneapolis, MN: University of Minnesota.  
<https://conservancy.umn.edu/bitstream/handle/11299/118720/39-1%262-Issue.pdf?sequence=1&isAllowed=y>
- Ross, M. B., Fazzalaro, J., Barone, K., & Kalinowski, J. (2016). State of Connecticut traffic stop data analysis and findings, 2014–15. New Britain, CT: Central Connecticut State University. <https://imrp.dpp.uconn.edu/wp-content/uploads/sites/3351/2021/09/May-2016.pdf>
- Smith, M. R. & Petrocelli, M. (2001). Racial profiling?: A multivariate analysis of police traffic stop data. *Police Quarterly*, 4, 4-27.
- Smith, M.R., Rojek, J., Tillyer, R., & Lloyd, C. (2017). *San Jose Police Department traffic and pedestrian stop study*. El Paso, TX: University of Texas at El Paso.  
[https://www2.sjpd.org/records/UTEP-SJPD\\_Traffic-Pedestrian\\_Stop\\_Study\\_2017.pdf](https://www2.sjpd.org/records/UTEP-SJPD_Traffic-Pedestrian_Stop_Study_2017.pdf)
- Smith, M.R., Tillyer, R., Lloyd, C., & Petrocelli, M. (2021). Benchmarking disparities in police stops: A comparative application of 2<sup>nd</sup> and 3<sup>rd</sup> generation techniques. *Justice Quarterly*, 38, 513-536. <https://doi.org/10.1080/07418825.2019.1660395>
- Smith, W. R., Tomaskovic-Devey, D., Zingraff, M. T., Mason, M., Warren, P. Y., & Wright, C. P. (2004). The North Carolina highway traffic study. Final report to the National Institute of Justice. <https://www.ncjrs.gov/pdffiles1/nij/grants/204021.pdf>.
- Taniguchi, T., Hendrix, J., Aagaard, B., Strom, K., Levin-Rector, A., & Zimmer, S. (2016). A test of racial disproportionality in traffic stops conducted by the Raleigh Police Department. Triangle Park, NC: RTI International. <https://www.rti.org/publication/test-racial-disproportionality-traffic-stops-conducted-raleigh-police-department>
- Tillyer, R., Engel, R. S., & Cherkauskas, J. C. (2010). Best Practices in Vehicle Stop Data Collection and Analysis. *Policing: Int'l J. Police Strat. & Mgmt.*, 33, 69-92.

Tomaskovic-Devey, D., Mason, M., & Zingraff, M. (2004). Looking for the driving while black phenomena: Conceptualizing racial bias processes and their associated distributions. *Police Quarterly*, 7, 3-29.

*United States v. Johnson*, 122 F. Supp. 3d 272 (M.D.N.C. 2013).

Vito, A.G., Griffin, V.W., Vito, G.F., & Higgins, G.E. (2020). “Does daylight matter”? An examination of racial bias in traffic stops by the police. *Policing: An International Journal*, 43. <https://doi.org/10.1108/PIJPSM-04-2020-0055>

Withrow, B. L. and Williams H. (2015). Proposing a benchmark based on vehicle collision data in racial profiling research. *Criminal Justice Review*, 40 (1), 449-469.

Worden, R. E., McLean, S. J., & Wheeler, A. P. (2012). Testing for racial profiling with the veil-of-darkness method. *Police Quarterly*, 15, 92-111.  
<https://doi.org/10.1177%2F1098611111433027>

Zingraff, M. T., Mason, H. M., Smith, W. R., Tomaskovic-Devey, D., Warren, P., McMurray, H. L., & Fenlon, C. R. (2000). Evaluating North Carolina State Highway Patrol data: Citations, warnings, and searches in 1998. Raleigh, NC: North Carolina State University, Center for Crime and Justice Research.